(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
28 March 2002 (28.03.2002)

PCT

(10) International Publication Number
WO 02/24056 A2

(51) International Patent Classification[7]: A61B

(21) International Application Number: PCT/US01/29713

(22) International Filing Date: 20 September 2001 (20.09.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/234,087 20 September 2000 (20.09.2000) US

(71) Applicant (*for all designated States except US*): **CASE WESTERN RESERVE UNIVERSITY** [US/US]; Office of Technology Management, 2109 Adelbert Road, Cleveland, OH 44106 (US).

(72) Inventors; and
(75) Inventors/Applicants (*for US only*): **MARKOWITZ, Sanford** [US/US]; 28950 Bolingbrook Road, Pepper Pike, OH 44124 (US). **GRADY, William** [US/US]; 3603 Trimble Road, Nashville, TN 37215 (US).

(74) Agent: **DOCHERTY, Pamela, A.**; Calfee, Halter & Griswold LLP, 800 Superior Avenue Suite 1400, Cleveland, OH 44114 (US).

(81) Designated States (*national*): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

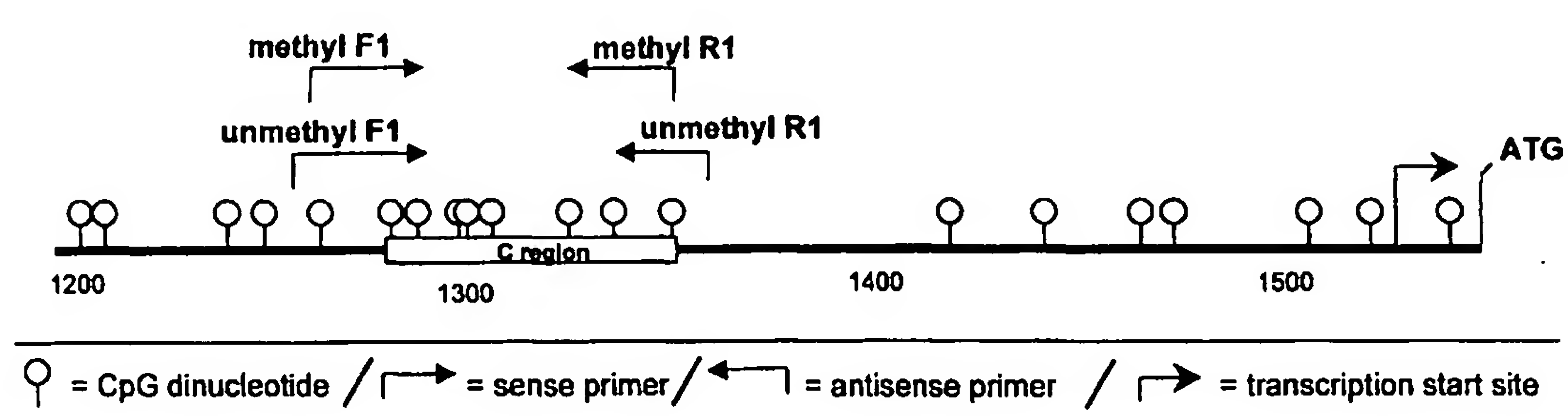
(84) Designated States (*regional*): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Declaration under Rule 4.17:
— *of inventorship (Rule 4.17(iv)) for US only*

Published:
— *without international search report and to be republished upon receipt of that report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: METHODS AND COMPOSITIONS FOR DETECTING CANCERS ASSOCIATED WITH METHYLATION OF *hMLH1* PROMOTER DNA



(57) Abstract: Methods are provided for detection of cancers associated with methylation of *hMLH1* promoter DNA in a subject. The method comprise assaying for the presence of methylated *hMLH1* promoter DNA in a bodily fluid from a subject. In one embodiment, the method comprises reacting DNA from the sample with a chemical compound that converts non-methylated cytosine bases, to a different nucleotide base. The compound-converted DNA is then amplified using a methylation-sensitive polymerase chain reaction (MSP) employing primers that amplify the compound-converted DNA template. The present invention also provides nucleotide primer sequences for use in the methylation-sensitive PCR assay.

WO 02/24056 A2

METHODS AND COMPOSITIONS FOR DETECTING CANCERS ASSOCIATED WITH METHYLATION OF *hMLH1* PROMOTER DNA

This invention was made, at least in part, with government support under National Institutes of Health Grant Nos: KO8 CA77676, RO1 CA67409 and RO1 CA72160. The U.S. government has certain rights in the invention.

CROSS REFERENCE TO RELATED APPLICATIONS

This invention claims priority to United States Provisional Patent Application Serial No.: 60/234,087, filed September 20, 2000.

BACKGROUND

In 2001, over 1.2 million new cases of human cancer will be diagnosed and over 0.5 million people will die from cancer (American Cancer Society estimate). Despite this, more people than ever are living with and surviving cancer. In 1997, for example, approximately 8.9 million living Americans had a history of cancer (National Cancer Institute estimate). People are more likely to survive cancer if the disease is diagnosed at an early stage of development, since treatment at that time is more likely to be successful. Early detection depends upon availability of high-quality methods. Such methods are also useful for determining patient prognosis,

selecting therapy, monitoring response to therapy and selecting patients for additional therapy. Consequently, there is a need for cancer diagnostic methods that are specific, accurate, minimally invasive, technically simple and inexpensive.

Colorectal cancer (i.e., cancer of the colon or rectum) is one particularly important type of human cancer. Colorectal cancer is the second most common cause of cancer mortality in adult Americans (Landis, et al., 1999, CA Cancer J Clin, 49:8-31). Approximately 40% of individuals with colorectal cancer die. In 2001, it is estimated that there will be 135,400 new cases of colorectal cancer (98,200 cases of colon and 37,200 cases of rectal cancer) and 56,700 deaths (48,000 colon cancer and 8,800 rectal cancer deaths) from the disease (American Cancer Society). As with other cancers, these rates can be decreased by improved methods for diagnosis.

Although methods for detecting colorectal cancer exist, the methods are not ideal. Digital rectal exams (i.e., manual probing of rectum by a physician), for example, although relatively inexpensive, are unpleasant and can be inaccurate. Fecal occult blood testing (i.e., detection of blood in stool) is nonspecific because blood in the stool has multiple causes. Colonoscopy and sigmoidoscopy (i.e., direct examination of the colon with a flexible viewing instrument) are both uncomfortable for the patient and expensive. Double-contrast barium enema (i.e., taking X-rays of barium-filled colon) is also an expensive procedure, usually performed by a radiologist.

Additional methods for detecting colorectal cancer require tissue from the tumor. In order to use these methods, growth of the tumor must have progressed far enough that the tumor is visible. A second disadvantage of these methods is that they require a tissue biopsy to be taken from the tumor.

Because of the disadvantages of existing methods for detecting colorectal cancer, new methods are needed.

SUMMARY OF THE INVENTION

In accordance with the present invention, new methods are provided for detection of cancers associated with methylation of *hMLH1* promoter DNA in a subject. In one aspect, the

method comprises assaying for the presence of methylated *hMLH1* promoter DNA in a bodily fluid from a subject. Preferably, the sample is blood, serum, plasma, a blood-derived fraction, stool, colonic effluent or urine. Preferably, such method comprises reacting DNA from the sample with a chemical compound that converts non-methylated cytosine bases (also called "conversion-sensitive" cytosines), but not methylated cytosine bases, to a different nucleotide base. In a preferred embodiment, the chemical compound is sodium bisulfite, which converts unmethylated cytosine bases to uracil. The compound-converted DNA is then amplified using a methylation-sensitive polymerase chain reaction (MSP) employing primers that amplify the compound-converted DNA template if cytosine bases within CpG dinucleotides of the DNA from the sample are methylated. Production of a PCR product indicates that the subject has a cancer associated with methylation of *hMLH1* promoter DNA.

In another aspect, the present invention provides a detection method for prognosis of a cancer in a subject known to have or suspected of having a cancer associated with methylation of *hMLH1* promoter DNA, referred to hereinafter as an "*hMLH1* methylation-associated cancer". Such method comprises assaying for the presence of methylated *hMLH1* promoter DNA in a bodily fluid from the subject. In certain cases, presence of the methylated *hMLH1* promoter DNA in the bodily fluid indicates that the subject is a good candidate for a particular therapy. In other cases, presence of the methylated *hMLH1* promoter DNA in the bodily fluid indicates that the *hMLH1* methylation associated cancer has a poor prognosis or the subject is a candidate for more aggressive therapy.

In another aspect, the present invention provides a method for monitoring over time the status of an *hMLH1* methylation-associated cancer in a subject. The method comprises assaying for the presence of methylated *hMLH1* promoter DNA in a bodily fluid taken from the subject at a first time and in a corresponding bodily fluid taken from the subject at a second time. Absence of methylated *hMLH1* promoter DNA from the bodily fluid taken at the first time and presence of methylated *hMLH1* promoter DNA in the bodily fluid taken at the second time indicates that the cancer is progressing. Presence of methylated *hMLH1* promoter DNA in the bodily fluid taken at the first time and absence of methylated *hMLH1* promoter DNA from the bodily fluid taken at the second time indicates that the cancer is regressing.

In another aspect the present invention provides a method for evaluating therapy in a subject suspected of having or having an *hMLH1* methylation-associated cancer. The method comprises assaying for the presence of methylated *hMLH1* promoter DNA in a bodily fluid taken from the subject prior to therapy and a corresponding bodily fluid taken from the subject during or following therapy. Loss of methylated *hMLH1* promoter DNA in the sample taken after or during therapy as compared to the presence of methylated *hMLH1* promoter DNA in the sample taken before therapy is indicative of a positive effect of the therapy on cancer regression in the treated subject

The present invention also provides nucleotide primer sequences for use in the methylation-sensitive PCR assay.

BRIEF DESCRIPTION OF THE FIGURES

Figure 1. Location of primers for *hMLH1* methylation-sensitive polymerase chain reaction (MSP) assay in relation to the *hMLH1* promoter C region identified by Deng *et al.* (Deng, et al., 1999, Cancer Res, 59:2029-33). Primers cover 5 CpG dinucleotides in this region and are designed to specifically amplify chemically-converted sequences arising from either methylated ("methyl") or unmethylated ("unmethyl") templates. "F" indicates forward primer and "R" indicates reverse primer. The numbering scheme is based on the GenBank sequence entry U83845 (Kane, et al., 1997, Cancer Res, 57:808-11).

Figure 2. Sequence level illustration of part of the *hMLH1* promoter (i.e. the Watson strand), showing location and sequence of primers. Primers are in bold type. The numbering scheme is based on the GenBank sequence entry U83845 (Kane, et al., 1997, Cancer Res, 57:808-11).

Figure 3. Double-stranded DNA sequence of the *hMLH1* promoter (GenBank accession number U83845) (Kane, et al., 1997, Cancer Res, 57:808-11).

Figure 4: Results of *hMLH1* MSP assay using primers that amplify methylated (M) or unmethylated (U) alleles specifically. M1 and M2 are DNA samples from cell lines V5 and RKO, respectively, previously shown to have only methylated *hMLH1* alleles; whereas, U1 and U2 are DNA samples from cell lines SW480 and V457, respectively, that have only unmethylated *hMLH1* alleles (Deng, et al., 1999, Cancer Res, 59:2029-33; Veigl, et al., 1998, Proc Natl Acad Sci U S A, 95:8698-702).

Figure 5: Results of *hMLH1* MSP from colon cancers and paired serum samples. (M) and (U), respectively, designate assays specific for amplification of the Methylated versus Unmethylated *hMLH1* promoter. The unmethylated alleles detected in tumor samples likely

represent contaminating normal tissue elements, and in serum samples likely arise from DNA released from circulating leukocytes.

DETAILED DESCRIPTION OF THE INVENTION

The present invention provides a method for detecting cancers in a subject that are associated with methylation of *hMLH1* promoter DNA in a bodily fluid from the subject. The preferred bodily fluid is blood, serum, plasma, blood-derived fraction, stool, colonic effluent or urine.

Definitions

Herein, "Watson strand" refers to one specific strand of a double-stranded DNA. Herein, the Watson strand of the *hMLH1* promoter is the same strand of DNA that contains the coding sequence of the *hMLH1* gene. The RNA that is transcribed from the *hMLH1* gene is identical in sequence to the coding sequence of the *hMLH1* gene, except that thymine (T) bases in the DNA coding strand sequence are replaced by uracil (U) bases in the RNA. The Watson strand of the DNA is the strand of the DNA that is conventionally shown when a single strand of a double-stranded DNA gene sequence is pictured.

Herein, "Crick strand" refers to one specific strand of a double-stranded DNA. Herein, the Crick strand of the *hMLH1* promoter is the same strand of DNA that contains the noncoding sequence of the *hMLH1* gene. The RNA that is transcribed from the *hMLH1* gene is complementary in sequence to the noncoding sequence of the *hMLH1* gene (U bases in the RNA are complementary to adenine bases (A) in the noncoding sequence of the DNA). The Crick strand of the DNA is the strand of the DNA that is, by convention, complementary to the single strand of DNA shown when a double-stranded DNA gene sequence is pictured.

Herein, "conversion-sensitive cytosine" refers to cytosine (C) bases in DNA that are unmethylated (i.e., are not 5-methylcytosine).

Herein, "compound-converted DNA" refers to DNA that has been treated or reacted with a chemical compound that converts unmethylated C bases in DNA to a different nucleotide base. Herein, one such compound is sodium bisulfite, which converts unmethylated C to U. If DNA that contains conversion-sensitive cytosine is treated with sodium bisulfite, the compound-converted DNA will contain U in place of C. If the DNA which is treated with sodium bisulfite

contains only methylcytosine, the compound-converted DNA will not contain uracil in place of the methylcytosine.

For example, the diagram below shows how sodium bisulfite treatment results in sequence changes within a DNA strand, dependent upon whether C bases in the DNA are methylated or not. Diagram I below shows a single-stranded DNA sequence. Two C bases within the sequence shown in I are part of CpG dinucleotides (i.e., the base immediately 3' of the C is G). These two C bases are shaded in diagram I. Four other C bases are not part of CpG dinucleotides (i.e., the base immediately 3' of the C is A, C or T). These four C bases are underlined in diagram I.

In the case that the two C bases in diagram I that are part of CpG dinucleotides (the shaded C bases) are methylated (i.e., are 5-methylcytosine bases), sodium bisulfite treatment of the DNA strand shown in diagram I will produce the DNA sequence shown in diagram II. Diagram II shows that only C bases that are unmethylated are converted to U by sodium bisulfite. C bases that are methylated (i.e., the two shaded C bases in diagram I) are not converted to U, nor to any other base, by sodium bisulfite.

I

5'-TAAAAACGAA<u>CC</u>AATAGGAAGAG<u>C</u>GGA<u>C</u>AGCG-3'

II

5'-TAAAAACGAA<u>UU</u>AATAGGAAGAG<u>U</u>GGA<u>U</u>AGCG-3'

III

5'-TAAAAA[illegible]GAA<u>UU</u>AATAGGAAGAG<u>U</u>GGA<u>U</u>AG[illegible]G-3'

In the case that the two C bases in diagram I that are part of CpG dinucleotides (the shaded C bases) are not methylated, sodium bisulfite treatment of the DNA strand shown in diagram I will produce the DNA sequence shown in diagram III. Diagram III shows that all C bases in diagram I are converted to U by sodium bisulfite. Only C bases that are methylated are protected from conversion to U by sodium bisulfite. Since all C bases in diagram I (C bases that

are part of CpG dinucleotides, and C bases that are not part of CpG dinucleotides) are unmethylated in this scenario, all C bases are converted to U, as shown in diagram III.

Herein, "compound-converted template sequence" refers to a region of the *hMLH1* promoter DNA that is rich in CpG dinucleotide sequences (i.e. 2 or more CpG dinucleotides), that has been treated with a chemical compound that converts unmethylated C bases in DNA to a different nucleotide base, and that is amplified by a methylation-sensitive polymerase chain reaction (MSP). Herein, the compound-converted template sequence is single-stranded and is either the Watson or Crick strand of the DNA.

Herein, "target sequence" refers to regions of the compound-converted template sequence from which the primers used in the MSP are derived. These target sequences are located at both ends of the compound-converted template sequence that is amplified by the MSP. Target sequences are between 8 and 50 nucleotides in length. Primers derived from the target sequences at both ends of a compound-converted template sequence will amplify the compound-converted template sequence in a MSP reaction (see below).

Herein, the "second target sequence" is at the 3' end of the compound-converted template sequence. The second target sequence contains at least one CpG dinucleotide (if C in the CpG dinucleotides in the DNA treated with sodium bisulfite is methylated) or at least one UpG dinucleotide (if C in the CpG dinucleotides in the DNA treated with sodium bisulfite is not methylated). The second target sequence also contains at least one U that is immediately 5' of an A, T or C base.

During the first round of the MSP, the reverse primer (see below) anneals to the second target sequence of the compound-converted template sequence. The reverse primer is complementary in sequence to the second target sequence of the compound-converted template sequence. After the reverse primer anneals, the 3' end of the reverse primer is extended in the 5' direction by the polymerase present in the MSP. Extension of the reverse primer by the polymerase is called herein, the "first round" of the MSP.

Herein, the "first target sequence" is at the 5' end of the compound-converted template sequence. The first target sequence contains at least one CpG dinucleotide (if C in the CpG

dinucleotides in the DNA treated with sodium bisulfite is methylated) or at least one UpG dinucleotide (if C in the CpG dinucleotides in the DNA treated with sodium bisulfite is not methylated). The first target sequence also contains at least one U that is immediately 5' of an A, T or C base.

During the second round of the MSP, the forward primer (see below) anneals to the 3' end of the extended reverse primer that is the product of the first round of the MSP. The forward primer is identical in sequence to the "first target sequence," except that U bases are replaced by T bases. Thus, the sequence of the forward primer is "substantially identical" to the first target sequence. After the forward primer anneals, the 3' end of the forward primer is extended in the 5' direction by the polymerase present in the MSP. Extension of the forward primer by the polymerase is called herein, the "second round" of the MSP. Addition cycles of both the first round step and the second round step of the MSP gives rise to the MSP product whose presence is determinative of methylation in the *hMLH1* promoter (see discussion below).

Herein, "methylation-sensitive PCR" (MSP) refers to a polymerase chain reaction in which amplification of the compound-converted template sequence is performed. Two sets of primers are designed for use in MSP. Each set of primers comprises a forward primer and a reverse primer. One set of primers, called methylation-specific primers (see below), will amplify the compound-converted template sequence if C bases in CpG dinucleotides within the *hMLH1* promoter DNA are methylated. Another set of primers, called unmethylation-specific primers (see below), will amplify the compound-converted template sequences if C bases in CpG dinucleotides within the *hMLH1* promoter DNA are not methylated.

Herein, "methylation-specific primers" are primers that direct amplification of a compound-converted template sequence where C bases within CpG dinucleotides of the first target sequence and the second target sequence are methylated.

Herein, "unmethylation-specific primers" are primers that direct amplification of a compound-converted template sequence where C bases within CpG dinucleotides of the first target sequence and the second target sequence are not methylated.

Herein, "reverse primer" refers to a primer for MSP that is located at the 3' end of the compound-converted template sequence and anneals to the second target sequence within the compound-converted template sequence. The reverse primer is complementary in sequence to the second target sequence of the compound-converted template sequence. The reverse primer is also identical to the reverse complement of the second target sequence. The sequence of the reverse primer contains C bases that are complementary to G bases within the second target sequence, G bases that are complementary to C bases, T bases that are complementary to A bases, and A bases that are complementary to both T and U bases in the second target sequence.

Herein, "forward primer" refers to a primer for MSP that is located at the 5' end of the compound-converted template sequence and anneals to the 3' end of the product of the first round of the MSP. The forward primer is identical in sequence to the first target sequence, except that U bases in the second target sequence are replaced by T bases in the forward primer. Thus, the sequence of the forward primer is "substantially identical" to the first target sequence.

Herein, "blood-derived fraction" refers to a component or components of whole blood. Whole blood comprises a liquid portion (i.e., plasma) and a solid portion (i.e., blood cells). The liquid and solid portions of blood are each comprised of multiple components; e.g., different proteins in plasma or different cell types in the solid portion. One of these components or a mixture of any of these components is a blood-derived fraction as long as such fraction is missing one or more components found in whole blood.

Genetic Changes in Colon Cancer

A number of human cancers are associated with genetic changes in DNA mismatch repair genes. DNA mismatch repair genes encode proteins that correct errors in DNA that are made when DNA polymerases insert an incorrect nucleotide base at a specific position in the DNA sequence during replication of the DNA. Defects in DNA mismatch repair genes result in cellular mutation rates at least 100-times higher than in cells with intact DNA mismatch repair genes (Modrich and Lahue, 1996, Annu Rev Biochem, 65:101-33; Thomas, et al., 1996, Mutat Res, 350:201-5).

One manifestation of the high mutation rate associated with defects in DNA mismatch repair genes is instability of DNA microsatellite sequences. Microsatellite sequences are comprised of nucleotide motifs 1-4 nucleotides in length that are repeated, possibly 100,000 times, in the human genome. Defects in DNA mismatch repair genes result in differences in the lengths of the nucleotide motif repeats in cancer cells as compared to normal cells. The process that leads to the repeat length differences is called microsatellite instability (MSI) and is found in a number of human cancer types.

A number of different DNA mismatch repair genes exist in humans. Germ-line mutation and inactivation of any of these commonly give rise to familial MSI colon cancers, known as hereditary nonpolyposis colon cancers (Marra and Boland, 1995, J Natl Cancer Inst, 87:1114-25; Lynch and Smyrk, 1996, Cancer, 78:1149-67). More common than the inherited type are colorectal cancers that arise sporadically (i.e., arise in an individual but are not inherited). In certain sporadic colorectal cancers, genetic changes in one particular DNA mismatch repair gene, *hMLH1*, are prominent.

Although alterations in the *hMLH1* gene are frequent in sporadic colorectal cancer, only a small percentage of these alterations are genetic mutations that change the DNA sequence of the *hMLH1* gene. Rather, the *hMLH1* genes in these sporadic cancers more often display epigenetic changes, specifically, changes in their DNA methylation patterns. In mammalian cells, DNA methylation comprises addition of a methyl group to the 5-carbon position of cytosine (C) nucleotide bases to form 5-methylcytosine (5mC) or methylcytosine. Only cytosines located 5' to guanines (G) in CpG dinucleotides are methylated in mammalian cells, but not all CpG dinucleotides are methylated. The pattern of methylation (i.e., whether a particular CpG dinucleotide is methylated or not) is relatively constant in cells (i.e., is maintained as cells divide), but can change under various circumstances. In such cases, the changed methylation pattern becomes relatively constant and is maintained.

One circumstance in which methylation patterns in cells can change is in certain human cancers. In such cancers, the methylation status (whether methylated or unmethylated) of certain CpG dinucleotides changes. In 10% of colon cancers (and 90% of MSI colon cancers), for example, there is methylation of C bases in CpG dinucleotides in the promoter region of the

hMLH1 gene. Ten to fifteen percent of gastric cancers also have methylation of C bases in CpG dinucleotides in the *hMLH1* promoter. Methylated C bases in CpG dinucleotides of *hMLH1* promoter DNA are also found in endometrial cancer, ovarian cancer and in precancerous colon adenomas.

Once established, the methylation of C bases within CpG dinucleotides in cancers is stable. These C bases are not methylated in non-cancerous cells (Kane, et al., 1997, Cancer Res, 57:808-11; Veigl, et al., 1998, Proc Natl Acad Sci U S A, 95:8698-702; Herman, et al., 1998, Proc Natl Acad Sci U S A, 95:6870-5). Methylation of C bases within CpG dinucleotides within nucleotides 1264 to 1354 of the *hMLH1* promoter (as numbered in GenBank sequence entry U83845; Figs. 2 and 3 of this document) are correlated with lack of expression of the *hMLH1* gene in colorectal cancer (Deng, et al., 1999, Cancer Res, 59:2029-33).

Detection of Changed Methylation in *hMLH1* Promoter DNA

The *hMLH1* promoter DNA, in its methylated state, is thus a cancer-specific modification that can serve as a target for detection using assay methods. One assay for detecting such methylated nucleotides is based on treatment of genomic DNA with a chemical compound which converts non-methylated C, but not methylated C (i.e., 5mC), to a different nucleotide base. One such compound is sodium bisulfite, which converts C, but not 5mC, to U. Methods for bisulfite treatment of DNA are known in the art (Herman, et al., 1996, Proc Natl Acad Sci U S A, 93:9821-6; Herman and Baylin, 1998, Current Protocols in Human Genetics, N. E. A. Dracopoli, ed., John Wiley & Sons, 2:10.6.1-10.6.10). When DNA that contains unmethylated C nucleotides is treated with sodium bisulfite to give compound-converted DNA, the sequence of that DNA is changed (C→U). Detection of U in the DNA is indicative of an unmethylated C.

The different nucleotide base (e.g., U) present in compound-converted DNA sequences can be detected in a variety of ways. One method of detecting U in compound-converted DNA sequences is a method called “methylation sensitive PCR” (MSP). In MSP, one set of primers (i.e., a set of primers comprises a forward and a reverse primer) amplifies the compound-converted template sequence if C bases in CpG dinucleotides within the *hMLH1* promoter DNA are methylated. This set of primers is called “methylation-specific primers.” Another set of primers amplifies the compound-converted template sequence if C bases in CpG dinucleotides

within the *hMLH1* promoter DNA are not methylated. This set of primers is called "unmethylation-specific primers."

Two separate PCR reactions are then run. Both reactions use the compound-converted DNA from bodily fluid. In one of the reactions, methylation-specific primers are used. In the case where C within CpG dinucleotides of the target sequence of the DNA are methylated, the methylation-specific primers will amplify the compound-converted template sequence in the presence of a polymerase and an MSP product will be produced. If C within CpG dinucleotides of the target sequence of the DNA are not methylated, the methylation-specific primers will not amplify the compound-converted template sequence in the presence of a polymerase and an MSP product will not be produced.

In the other reaction, unmethylation-specific primers are used. In the case where C within CpG dinucleotides of the target sequence of the DNA are unmethylated, the unmethylation specific primers will amplify the compound-converted template sequence in the presence of a polymerase and an MSP product will be produced. If C within CpG dinucleotides of the target sequence of the DNA are methylated, the unmethylation-specific primers will not amplify the compound-converted template sequence in the presence of a polymerase and an MSP product will not be produced.

Selection of Primers for MSP

Primers are derived from target sequences at the ends of the compound-converted *hMLH1* promoter template sequence. Herein, "derived from" means that the sequences of the primers are chosen such that the primers amplify the compound-converted template sequence, which includes the target sequences, in an MSP reaction. The second target sequence is at the 3' end of the compound-converted template sequence. The first target sequence is at the 5' end of the compound-converted template sequence. Target sequences are between 8 and 50 nucleotides in length, preferably from 25 to 32 nucleotides in length. The target sequences contain at least one CpG dinucleotide, preferably two or more CpG dinucleotides (if C in the CpG dinucleotides in the DNA treated with sodium bisulfite is methylated), or at least one UpG dinucleotide, preferably two or more UpG dinucleotides (if C in the CpG dinucleotides in the DNA treated with sodium bisulfite is not methylated). Preferably, one of the CpG or UpG dinucleotides is

located at the 5' end of the second target sequence. Preferably one of the CpG or UpG dinucleotides is located at the 3' end of the first target sequence. The target sequences also contain at least one U that is immediately 5' of an A, T or C base.

Each primer comprises a single-stranded DNA fragment which is at least 8 nucleotides in length. Preferably, the primers are less than 50 nucleotides in length, more preferably from 15 to 35 nucleotides in length. The sequences of the primers are derived from the target sequences. The reverse primer is complementary in sequence to the second target sequence of the compound-converted template sequence. The reverse primer is also identical to the reverse complement of the second target sequence. The sequence of the reverse primer contains C bases that are complementary to G bases within the second target sequence, G bases that are complementary to C bases, T bases that are complementary to A bases, and A bases that are complementary to both T and U bases in the second target sequence. The forward primer is identical in sequence to the first target sequence, except that U bases in the first target sequence are replaced by T bases in the forward primer. Thus, the sequence of the forward primer is substantially "identical" to the first target sequence.

Because the compound-converted template sequence can be either the Watson strand or the Crick strand of the double-stranded DNA that is treated with sodium bisulfite, the sequences of the primers is dependent upon whether the Watson or Crick compound-converted template sequence is chosen to be amplified in the MSP. Either the Watson or Crick strand can be chosen to be amplified.

One set of primer sequences (a set of primers comprises a forward primer and a reverse primer), called methylation-specific primers, is chosen such that the primers amplify the compound-converted template sequence when the compound-converted template sequence contains one or more CpG dinucleotides. Another set of primer sequences, called unmethylation-specific primers, is chosen such that the primers amplify the compound-converted template sequence when the compound-converted template sequence contains one or more UpG dinucleotides.

The compound-converted template sequence, and therefore the product of the MSP reaction, can be between 20 to 2000 nucleotides in length, preferably between 50 to 500

nucleotides in length, more preferably between 80 to 150 nucleotides in length. Preferably, the methylation-specific primers result in an MSP product of a different length than the MSP product produced by the unmethylation-specific primers.

Samples and Isolation of DNA Therefrom

DNA which is used as the template in the MSP is obtained from a subject suspected of or known to have, or to have had, an *hMLH1* methylation-associated cancer. Alternatively, the template DNA is from a subject who is undergoing a routine screening and who is not necessarily suspected of having an *hMLH1* methylation-associated cancer. This DNA is obtained from a bodily fluid. Examples of preferred bodily fluids are blood, serum, plasma, a blood-derived fraction, stool, colonic effluent or urine. Other body fluids can also be used. Because they can be easily obtained from the subject and can be used to screen for multiple diseases, blood or blood-derived fractions are especially useful. It has been shown that DNA alterations in colorectal cancer patients can be detected in the blood of subjects (Hibi, et al., 1998, Cancer Res, 58:1405-7).

Blood-derived fractions can comprise blood, serum, plasma, or other fractions. For example, a cellular fraction can be prepared as a "buffy coat" (i.e., leukocyte-enriched blood portion) by centrifuging 5 ml of whole blood for 10 min at 800 times gravity at room temperature. Red blood cells sediment most rapidly and are present as the bottom-most fraction in the centrifuge tube. The buffy coat is present as a thin creamy white colored layer on top of the red blood cells. The plasma portion of the blood forms a layer above the buffy coat. Fractions from blood can also be isolated in a variety of other ways. One method is by taking a fraction or fractions from a gradient used in centrifugation to enrich for a specific size or density of cells.

DNA is then isolated from samples from the bodily fluids. Procedures for isolation of DNA from such samples are well known to those skilled in the art. Commonly, such DNA isolation procedures comprise lysis of any cells present in the samples using detergents, for example. After cell lysis, proteins are commonly removed from the DNA using various proteases. RNA is removed using RNase. The DNA is then commonly extracted with phenol, precipitated in alcohol and dissolved in an aqueous solution.

Methylation-Sensitive PCR (MSP)

Isolated DNA is reacted with a compound that converts unmethylated cytosine (C) bases in the DNA to another base. Preferably, the compound is sodium bisulfite, which converts unmethylated C to uracil (U). Portions of the compound-converted DNA is used in two different MSP reactions with different primers. In each MSP reaction, the primers comprise a forward primer and a reverse primer. In one of the MSP reactions, methylation-specific primers are used. In another of the MSP reactions, unmethylation-specific primers are used. Conditions for the MSP reactions are well known to those skilled in the art and are described in more detail in Example 5.

A variety of methods can be used to determine if an MSP product has been produced in a reaction assay. One way to determine if an MSP product has been produced in the reaction is to analyze a portion of the reaction by agarose gel electrophoresis. For example, a horizontal agarose gel of from 0.6 to 2.0% agarose is made and a portion of the MSP reaction mixture is electrophoresed through the agarose gel. After electrophoresis, the agarose gel is stained with ethidium bromide. MSP products are visible when the gel is viewed during illumination with ultraviolet light. By comparison to standardized size markers, it is determined if the MSP product is of the correct expected size.

Other methods can be used to determine whether a product is made in an MSP reaction. One such method is called "real-time PCR." Real-time PCR utilizes a thermal cycler (i.e., an instrument that provides the temperature changes necessary for the PCR reaction to occur) that incorporates a fluorimeter (i.e. an instrument that measures fluorescence). The real-time PCR reaction mixture also contains a reagent whose incorporation into a product can be quantified and whose quantification is indicative of copy number of that sequence in the template. One such reagent is a fluorescent dye, called SYBR Green I (Molecular Probes, Inc.; Eugene, Oregon) that preferentially binds double-stranded DNA and whose fluorescence is greatly enhanced by binding of double-stranded DNA. When a PCR reaction is performed in the presence of SYBR Green I, resulting DNA products bind SYBR Green I and fluorescence. The fluorescence is detected and quantified by the fluorimeter. Such technique is particularly useful for quantification of the amount of template in PCR reaction.

Other Methods for Detecting Methylated *hMLH1* Promoter DNA Sequences

A modification of MSP is that primers for use in MSP are designed and chosen such that products of the MSP reaction are susceptible to digestion by restriction endonucleases, dependent upon whether the compound-converted template sequence contains CpG dinucleotides or UpG dinucleotides.

Another technique referred to as MS-SnuPE, uses compound-converted DNA as a template in a primer extension reaction wherein the primers used produce a product, dependent upon whether the compound-converted template contains CpG dinucleotides or UpG dinucleotides.

Another technique is to directly sequence the product resulting from an MSP reaction to determine if the compound-converted template sequence contained CpG dinucleotides or UpG dinucleotides.

Another technique for detecting methylated *hMLH1* promoter DNA sequences uses restriction endonucleases called "methylation-sensitive" restriction endonucleases to treat the DNA isolated from the bodily fluid of a subject. The restriction endonuclease-treated DNA is then used as a template in a PCR reaction. Herein, methylation-sensitive restriction endonucleases recognize and cleave a specific sequence within the DNA if C bases within the recognition sequence are not methylated. If C bases within the recognition sequence of the restriction endonuclease are methylated, the DNA will not be cleaved. One such methylation-sensitive restriction endonuclease is Hpa II. Other such methylation-sensitive restriction endonucleases exist and are known in the art. In this technique, a recognition sequence for a methylation-sensitive restriction endonuclease is located within the template DNA, at a position between the forward and reverse primers used for the PCR reaction. In the case that a C base within the methylation-sensitive restriction endonuclease recognition sequence is not methylated, the endonuclease will cleave the DNA template and a PCR product will not be formed when the DNA is used as a template in the PCR reaction. In the case that a C base within the methylation-sensitive restriction endonuclease recognition sequence is methylated, the endonuclease will not cleave the DNA template and a PCR product will be formed when the DNA is used as a template in the PCR reaction. Therefore, methylation of C bases can be determined by the absence or

presence of a PCR product (Kane, et al., 1997, Cancer Res, 57:808-11). No sodium bisulfite is used in this technique.

Subjects from which Samples are Obtained

An important aspect of the invention is that DNA that is assayed in the MSP comes from human subjects that are suspected of having, has, or is known to have had a cancer associated with methylation of *hMLH1* promoter DNA. Alternatively, the DNA comes from a subject who is undergoing routine screening and who is not necessarily suspected of having an *hMLH1* associated cancer.

Detection of methylation of *hMLH1* promoter DNA in samples from subjects that are not known to have had a cancer associated with methylation of *hMLH1* promoter DNA can aid in diagnosis of such a cancer in the subject. The MSP assay of this invention can be used by itself, or in combination with other various assays, to improve the sensitivity and/or specificity of detection of the *hMLH1* methylation associated cancer. Preferably, such detection is made at an early stage in the development of cancer, so that treatment is more likely to be effective.

In addition to diagnosis, detection of *hMLH1* promoter DNA methylation in a sample from a subject not known to have had a cancer, can be prognostic for the subject (i.e., indicate the probable course of the disease). For example, cancers having a high probability of a favorable outcome in the subject may be associated with methylated *hMLH1* promoter DNA. Alternatively, the presence of methylated *hMLH1* promoter DNA in certain bodily fluids, such as blood, can indicate that the prognosis is poor. The prognostic potential of the method of the present invention may be particularly high if detection of *hMLH1* promoter methylation is indicative of a particular stage of the cancer or is associated with other characteristics of a cancer wherein the other characteristics are prognostic.

Detection of *hMLH1* promoter DNA methylation in samples from subjects can also be used to select a particular therapy or therapies which are particularly effective against the cancer in the subject, or to exclude therapies that are not likely to be effective. For example, colon cancer patients receiving adjuvant chemotherapy with 5-Fluorouracil had a much better survival rate if their tumors were MSI colon cancers (Watanabe, et al., 2001, N Engl J Med, 344:1196-

206). Ninety percent of MSI colon cancers have methylated *hMLH1* promoter DNA. Information of this type could also result in improved therapy design and, potentially, improved survival for the patient.

Detection of methylation of *hMLH1* promoter DNA in samples from subjects that are known to have, or to have had, a cancer associated with methylation of *hMLH1* promoter DNA is also useful. For example, the present methods can be used to identify subjects for whom therapy is or is not working. One or more bodily fluids are taken from the subject prior to and following therapy and assayed for the presence of methylated *hMLH1* promoter DNA. A finding that the methylated *hMLH1* promoter DNA is present in the sample taken prior to therapy and absent after therapy would indicate that the therapy is effective and need not be altered. In those cases where the methylated *hMLH1* promoter DNA is present in the sample taken before therapy and in the sample taken after therapy, it may be desirable to alter the therapy to increase the likelihood that the cancer will be eradicated in the subject. Thus, the present method could obviate the need to perform more invasive procedures which are used to determine a patient's response to therapy.

Cancers frequently recur following therapy in patients with advanced cancers. In this and other instances, the present detection methods are useful for monitoring over time the status of an *hMLH1* methylation associated cancer. For subjects in which cancer is progressing, methylated *hMLH1* promoter DNA may be absent from some or all bodily fluids when the first sample is taken and then appear in one or more bodily fluids when the second sample is taken. For subjects in which cancer is regressing, methylated *hMLH1* promoter DNA may be present in one or a number of bodily fluids when the first sample is taken and then be absent in some or all of these bodily fluids when the second sample is taken.

EXAMPLES

The following examples are for purposes of illustration only and are not intended to limit the scope of the invention as defined in the claims which are appended hereto.

Example 1 - Selection of Primers

To attain maximal sensitivity and specificity, primers for *hMLH1* MSP were designed that, following bisulfite conversion, selectively amplified compound-converted templates derived from either methylated (i.e., containing CpG dinucletides) or unmethylated (i.e., containing UpG dinucletides) versions of sequences contained within base pairs 1264 to 1354 of the *hMLH1* promoter DNA of GenBank entry U83845 (Fig. 2), a region whose methylation status has been demonstrated to invariably correlate with *hMLH1* expression in colon cancer (Deng, et al., 1999, Cancer Res, 59:2029-33). The primers annealed to a region of the *hMLH1* promoter sequence that contained a total of five CpG dinucleotides with at least one CpG dinucleotide placed at the 3' end of each primer to maximize discrimination between methylated and umethylated alleles after sodium bisulfite treatment. (Fig. 2.). To further enhance the discrimination between products amplified by the primer sets that amplified compound-converted templated sequences derived from methylated versus unmethylated *hMLH1* promoter DNA, the primers were designed to generate MSP products that varied in size by 11 bp (102 bp vs. 91 bp). The short size of these MSP products also increased the utility of these primers for determining the methylation status of the *hMLH1* promoter in samples of paraffin-embedded, formalin-fixed tumor tissues. The methylated and unmethylated primer pair sequences were as shown in Table 1 below. The primer pairs were purchased from Sigma-Genosys (The Woodlands, Texas).

Table 1. Primers for amplification of the Watson strand in MSP PCR

Primer name	Sequence (5' → 3')	SEQ. ID. NO.
Methyl specific forward primer 1	AACGAATTAATAGGAAGAGCGGATAGCG	2
Methyl specific reverse primer 1	CGTCCCTCCCTAAAACGACTACTACCC	3
Unmethyl specific forward primer 1	AAAAATGAATTAATAGGAAGAGTGGATAGTG	4
Unmethyl specific forward primer 2	TAAAAATGAATTAATAGGAAGAGTGGATAGTG	5
Unmethyl specific reverse primer 1	AATCTCTTCATCCCTCCCTAAAACA	6

Example 2 - Cell Lines and DNA Isolation

The cell lines RKO, SW480, Vaco-5, and Vaco-457 were used. RKO cells were obtained from M. Brattain at the Univesity of Texas Health Science Center, San Antonio, Texas. SW480 cells were obtained from the American Type Culture Collection, 10801 University Boulevard, Manassas, VA 20110-2209. The Vaco cell lines were established from colon cancers as described previously (Markowitz, et al., 1995, Science, 268:1336-8; Willson, et al., 1987, Cancer Res, 47:2704-13). Cells were cultured in MEM supplemented with gentamycin, L-glutamine, nonessential amino acids and sodium selenite. For Vaco-5 cells, growth medium was supplemented with 8% calf serum. For RKO, SW480 and Vaco-457 cells, growth medium was supplemented with 2% fetal bovine serum (HyClone), 10 mg/liter bovine insulin, 2 mg/liter human transferin and 1 mg/liter hydrocortisone. Monolayer cultures were grown at 37°C in a 5% CO_2 atmosphere. Genomic DNA was obtained by lysing cells in proteinase K buffer [10 mM

Tris (pH 7.4)/10 mM EDTA/0.4% SDS/150 mM NaCl/100 μg/ml proteinase K] and extracting with phenol/chloroform/isoamyl alcohol (25:24:1).

Example 3 - Serum as Bodily Fluid

Serum samples were prepared from whole blood immediately after collection of the blood. Serum samples were prepared by centrifuging the sample at 2700 x *g* for 15 minutes and aspirating the serum. The serum was then stored at -80°C until use. DNA was extracted from the serum by incubating 1 ml serum in a buffered extraction solution (140 mM Tris/140 mM EDTA/0.57% SDS) with proteinase K (500 μg) overnight at 50°C in a shaking water bath. The sample was then subject to phenol:chloroform/chloroform DNA extraction twice, chloroform extraction, and ethanol precipitation. The precipitated DNA was resuspended in 25 μl distilled, deionized water and stored at -20°C. Approximately half of the sample was employed in the subsequent MSP assay.

Example 4 - Bisulfite Treatment of DNA

DNA (1 μg) was treated with sodium bisulfite by dilution in 50 μl of distilled water and denaturation with 0.2 M sodium hydroxide for 15 minutes at 37°C. Thirty μl of 10 mM hydroquinone (Sigma) and 520 μl of 3 M sodium bisulfite (Sigma) at pH 5.0 were then added, and the samples were incubated at 53°C for 18-20 hours. After treatment, the DNA was purified using the Wizard DNA Clean-Up System (Promega) following the manufacturer's protocol. The DNA was then desulphonated with 0.3 M sodium hydroxide for 10 minutes at room temperature, neutralized with 17 μl 10 M NH_4OAc, and then precipitated in 100% ethanol overnight at -80°C. The samples were resuspended in distilled water at a final concentration of approximately 50 ng/μl and stored at -80°C for up to 8 weeks. A set of 4 known methylated and unmethylated control DNA samples were included in each round of bisulfite treatment.

Example 5 - Methylation Specific PCR (MSP)

The compound-converted DNA was then subject to methylation specific PCR (MSP) using primer pairs as described in Example 1. The specific primers used were: methyl specific, SEQ. ID. NO. 2 and SEQ. ID. NO. 3; unmethyl specific, SEQ. ID. NO. 5 and SEQ. ID. NO. 6.

The PCR was performed in 20-μl reaction volumes containing 1X PCR buffer II (Perkin-Elmer), 1.5mM $MgCl_2$, 200μM dNTP mixture (Perkin-Elmer), 1μM of each primer, 100-150 ng modified DNA, and 1.5 units Amplitaq gold DNA polymerase (Perkin-Elmer). The thermocycler conditions were as follows: methyl-specific-- 95°C x 10', (92°C x 30", 62°C x 30", 70°C x 30") x 39 cycles, final extension of 70°C x 7'; and unmethyl specific-- 95°C x 10', (92°C x 30", 57°C x 30", 70°C x 30") x 39 cycles, final extension of 70°C x 7'. The PCR products were subject to gel electrophoresis through a 2.5% agarose gel, stained with ethidium bromide, and then visualized with UV illumination using a digital imaging system (Alpha Inotech).

Example 6 - Testing of Primers in MSP

The primers of Example 1 were tested for their specificity in amplifying compound-converted template sequences derived from methylated or unmethylated *hMLH1* promoter DNA. The DNA of four cell lines (Example 2) that had been characterized as, respectively, demonstrating an unmethylated, active or a methylated inactive *hMLH1* promoter was reacted with bisulfite (Example 4) and MSP was performed as in Example 5. DNA from cell lines V5 and RKO are known to have only methylated *hMLH1* alleles (Deng, et al., 1999, Cancer Res, 59:2029-33; Veigl, et al., 1998, Proc Natl Acad Sci U S A, 95:8698-702). DNA from cell lines SW480 and V457 are known to have only unmethylated *hMLH1* alleles (Deng, et al., 1999, Cancer Res, 59:2029-33; Veigl, et al., 1998, Proc Natl Acad Sci U S A, 95:8698-702).

As shown in Fig. 4, a PCR product was obtained with the methylation-specific primer only from those cell lines known to have a methylated *hMLH1* promoter.

Example 7 - Sensitivity Test

To determine the sensitivity of the methylation specific PCR assay for detecting compound-converted template sequences from methylated *hMLH1*, DNA from cell lines with methylated *hMLH1* alleles was mixed with DNA from cell lines bearing unmethylated *hMLH1* alleles. These mixtures were treated with sodium bisulfite (Example 4) and then subjected to the MSP assay (Example 5).

The results of this study showed that a minimum of 1.0% of methylated DNA could be detected with this assay. Based on the amount of input DNA used in these reactions, the assay

successfully detected the methylated *hMLH1* promoter in 1.5 ng of total genomic DNA, which is equivalent to the amount of DNA present in 100-200 cells.

Example 8 - MSP of DNA from Tumors

Primary colon cancer samples were obtained from nineteen patients. The MSI status of these colon cancers had been previously characterized. Of ten MSI colon cancers, ninety percent (n=9/10) demonstrated methylation of the *hMLH1* promoter. In contrast, 0 of 9 microsatellite stable (MSS) colon cancer demonstrated *hMLH1* promoter methylation. (See Table 2). Prior to the surgical resection of the tumors, matched serum samples were obtained from the patients.

Using the *hMLH1* promoter MSP assay (Example 5), the DNAs extracted from the preoperative serum samples that matched these same 19 colon cancers were tested, after they had been treated with bisulfite (Example 4). Of nine individuals whose MSI colon cancers tested positive for *hMLH1* promoter methylation, 33% (n=3/9) also demonstrated a positive assay for *hMLH1* promoter methylation in DNA from their preoperative serum samples (See Table 2 and Figure 5). A positive serum result was obtained in 2 of 2 individuals found at surgery to have distant metastases (stage D); in 1 of 3 individuals found to have nodal metastases (stage C), and in none of 4 individuals whose tumors proved to be confined to the colon (stage B). None of ten preoperative serum samples tested positive in the cases of individuals whose tumors lacked *hMLH1* promoter methylation (Table 2). Individuals whose tumors lacked *hMLH1* promoter methylation were similar to those whose tumors demonstrated *hMLH1* promoter methylation with regards to stage, location of the primary tumors, and age of the patients at diagnosis. Thus, in this sample set, assay of serum DNA for *hMLH1* promoter methylation demonstrated a sensitivity of 33% and a specificity of 100% for the detection of colon cancers bearing this molecular alteration.

Table 2. Results of *hMLH1* MSP in tumors and serum of colon cancer patients

Description of patient and tumor characteristics compared with results of *hMLH1* MS-PCR analysis of matched tumor and serum samples.

Patient no.	*hMLH1* Tumor DNA/Serum DNA[a]	MSI status	Stage	Grade
1	M/M	MSI	D	Poor
2	M/M	MSI	D	Moderate
3	M/M	MSI	C	Poor
4	M/U	MSI	C	Moderate
5	M/U	MSI	C	Moderate
6	M/U	MSI	B2	Moderate
7	M/U	MSI	B	Moderate
8	M/U	MSI	B	Moderate
9	M/U	MSI	A	Moderate
10	U/U	MSS	D	Moderate
11	U/U	MSS	D	Moderate
12	U/U	MSS	D	Poor
13	U/U	MSS	C	Moderate
14	U/U	MSS	B	Moderate
15	U/U	MSS	B	Moderate
16	U/U	MSS	B	Moderate
17	U/U	MSI	B	Moderate
18	U/U	MSS	B	Moderate

19	U/U	MSS	HGD[b]	HGD

[a]M, methylated *hMLH1* promoter; U, unmethylated *hMLH1* promoter

[b]HGD, high-grade dysplasia in adenoma

CLAIMS

What is claimed is:

1. A method for detecting the presence of a cancer associated with methylation of *hMLH1* promoter DNA in a subject, comprising:

assaying for the presence of methylated *hMLH1* promoter DNA in a bodily fluid from the subject.

2. The method of claim 1, wherein the bodily fluid is selected from the group consisting of blood, serum, plasma, a blood-derived fraction, stool, urine, and a colonic effluent.

3. The method of claim 1 wherein the bodily fluid is from a subject suspected of having or known to have colorectal cancer.

4. The method of claim 1 wherein the bodily fluid is from a subject suspected of having or known to have gastric cancer.

5. The method of claim 1 wherein the bodily fluid is from a subject suspected of having or known to have endometrial cancer.

6. The method of claim 1 wherein the bodily fluid is from a subject suspected of having a precancerous colon adenoma of the class in which *hMLH1* promoter is methylated.

7. The method of claim 1 wherein the bodily fluid is from a subject suspected of having or known to have ovarian cancer.

8. The method of claim 1 wherein the assay comprises the steps of:

a) treating DNA which has been extracted from said bodily fluid with a compound that converts non-methylated cytosine bases in said DNA to a different base;

c) amplifying said treated DNA by a polymerase chain reaction which employs a forward primer and a reverse primer designed to amplify a compound-converted template in a CpG rich region of the Watson strand of *hMLH1* promoter DNA or a compound-converted template in a CpG rich region of the Crick strand of *hMLH1* promoter DNA,

wherein each of said primers comprises at least 8 consecutive nucleotides derived from a target sequence in the compound-converted template;

wherein the compound-converted template of the Watson strand is amplified using a forward primer comprising a sequence which is substantially identical to a first target sequence said first target sequence being located at the 5' end of the compound-converted template of the Watson strand, and a reverse primer comprising a sequence which is identical to the reverse complement of a second target sequence, said second target sequence being located at the 3' end of the compound-converted template of the Watson strand;

wherein the compound-converted template of the Crick strand is amplified using a forward primer comprising a sequence which is substantially identical to a first target sequence said first target sequence being located at the 5' end of the compound-converted template of the Crick strand, and a reverse primer comprising a sequence which is identical to the reverse complement of a second target sequence, said second target sequence being located at the 3' end of the compound-converted template of the Crick; and

wherein the cytosines in said CpG rich region are not methylated in normal subjects.

9. The method of claim 8 wherein the CpG region is located between base pairs 1264 and 1354 within the *hMLH1* promoter DNA of Genbank Entry U83845.

10. The method of claim 8 wherein the compound-converted template on the Watson strand of the *hMLH1* promoter DNA is amplified using a forward primer comprising a sequence which is substantially identical to a portion or all of the compound-converted sequence extending from and including nucleotide 1264 through nucleotide 1291 of SEQ ID NO.1; and a reverse primer comprising a sequence which is identical to the reverse complement of a portion or all of the compound-converted sequence extending from and including nucleotide 1328 through nucleotide 1354 of SEQ ID NO. 1

11. The method of claim 10 wherein the forward primer comprises a sequence which comprises at least 8 consecutive nucleotides in the sequence 5'AACGAATTAATAGGAAGAGCGGATAGCG-3', SEQ ID NO. 2, and the reverse primer

comprises at least 8 consecutive nucleotides in the sequence 5'-CGTCCCTCCCTAAAACGACTACTACCC-3', SEQ ID NO. 3.

12. The method of claim 8 wherein the forward primer, the reverse primer, or both have a CpG at the 5' or 3' end thereof.

13. The method of claim 11 wherein the forward primer comprises at least 15 consecutive nucleotides in SEQ ID NO. 2 and the reverse primer comprises at least 15 consecutive nucleotides in SEQ ID NO. 3.

14. The method of claim 11 wherein the reverse primer and forward primer are less than 50 nucleotides in length, and

wherein the sequences derived from SEQ ID NO. 2 and SEQ ID NO. 3 are at the 3' end of the forward primer and the reverse primer, respectively.

15. The method of claim 8 wherein the compound-converted template on the Crick strand of the *hMLH1* promoter DNA is amplified using a forward primer comprising a sequence which is substantially identical to a portion or all of the compound-converted sequence extending from and including nucleotide 1354 through nucleotide 1326 of the complement to SEQ ID NO.1; and a reverse primer comprising a sequence which is identical to the reverse complement of a portion or all of the compound-converted sequence extending from and including nucleotide 1291 through nucleotide 1264 of the reverse complement of SEQ ID NO. 1

16. The method of claim 15 wherein the forward primer comprises at least 8 consecutive nucleotides in the sequence 5'-CGTTTTTTTTGAAGCGGTTATTGTTTGT-3', SEQ ID NO. 7,

and wherein the reverse primer comprises at least 8 consecutive nucleotides in the sequence 5'-AACGAACCAATAAAAAAACAAACAACG-3', SEQ ID NO. 8,

17. The method of claim 1 wherein the assay comprises the steps of:

a) treating DNA which has been extracted from said bodily fluid with sodium bisulfite to provide a bisulfite-converted DNA sample in which the non-methylated cytosine bases in said DNA sample are converted to uracil bases;

b) amplifying said sodium bisulfite converted DNA by a polymerase chain reaction which employs a forward primer and a reverse primer derived from the bisulfite converted Watson strand or the bisulfite converted Crick strand of a CpG rich region of *hMLH1* promoter DNA,

wherein each of said primers comprises at least 8 consecutive nucleotides derived from a target sequence in the bisulfite-converted template;

wherein the bisulfite-converted template of the Watson strand is amplified using a forward primer comprising a sequence which is substantially identical to a first target sequence said first target sequence being located at the 5' end of the bisulfite-converted template of the Watson strand, and a reverse primer comprising a sequence which is identical to the reverse complement of a second target sequence, said second target sequence being located at the 3' end of the bisulfite-converted template of the Watson strand;

wherein the bisulfite-converted template of the Crick strand is amplified using a forward primer comprising a sequence which is substantially identical to a first target sequence said first target sequence being located at the 5' end of the bisulfite-converted template of the Crick strand, and a reverse primer comprising a sequence which is identical to the reverse complement of a second target sequence, said second target sequence being located at the 3' end of the bisulfite-converted template of the Crick; and

wherein the cytosines in said CpG rich region are not methylated in normal subjects.

18. The method of claim 17 wherein each primer comprises a CpG dinucleotide.

19. The method of claim 17 wherein the forward primer, or the reverse primer, or both comprise at least two CpG dinucleotides.

20. The method of claim 17 wherein the forward primer comprises at least two thymines which correspond to two uracils that are immediately 5' of an A, C, or T base in the bisulfite-converted template sequence.

21. The method of claim 17 wherein the reverse primer comprises at least two adenines which are complementary to two uracils that are immediately 5' of an A, C, or T base in the bisulfite-converted template sequence

22. A method for monitoring over time a cancer associated with methylation of *hMLH1* promoter DNA in a subject; comprising

a) assaying for the presence of methylated *hMLH1* promoter DNA in a bodily fluid taken from the subject at a first time, said bodily sample selected from the group consisting of blood, serum, plasma, a blood-derived fraction, stool, and a colonic effluent, and

b) assaying for the presence of methylated *hMLH1* promoter DNA in a corresponding bodily fluid taken from the subject at a later time;

wherein absence of methylated *hMLH1* promoter DNA in the bodily fluid taken at a later time and the presence of methylated *hMLH1* promoter DNA in the bodily fluid taken at the first time is indicative of cancer regression and wherein presence of methylated *hMLH1* promter DNA in the bodily fluid taken at a later time and absence of methylated *hMLH1* promoter DNA in the bodily fluid taken at the first time is indicative of cancer progression.

23. The method of claim 22 wherein the bodily fluid is selected from the group consisting of blood, serum, plasma, a blood-derived fraction, stool, urine, and a colonic effluent.

24. A method for monitoring response to therapy of a cancer associated with methylation of *hMLH1* promoter DNA in a subject, comprising:

a) assaying for the presence of methylated *hMLH1* promoter DNA in a bodily sample taken from the subject prior to therapy, said bodily sample selected from the group consisting of blood, serum, plasma, a blood-derived fraction, stool, and a colonic effluent, and

b) assaying for the presence of methylated *hMLH1* promoter DNA in a corresponding bodily sample taken from the subject following or during therapy;

wherein loss of methylated *hMLH1* promoter DNA in the sample taken after or during therapy as compared to the presence of methylated *hMLH1* promoter DNA in the sample taken before therapy is indicative of a positive effect of said therapy on cancer regression in the treated subject.

25. The method of claim 24 wherein the sample is from a subject who has undergone treatment for colon cancer.

26. The method of claim 1 wherein the sample is from a subject who has undergone treatment for said cancer.

27. A detection method for prognosis of a cancer in a subject known to have or suspected of having a cancer associated with methylation of *hMLH1* promoter DNA, comprising:

assaying for the presence of methylated *hMLH1* promoter DNA in a bodily fluid from the subject;

wherein presence of the methylated *hMLH1* promoter DNA in the bodily fluid indicates said subject is a good candidate for a particular therapy.

28. A detection method for prognosis of a cancer in a subject known to have or suspected of having a cancer associated with methylation of *hMLH1* promoter DNA, comprising:

assaying for the presence of methylated *hMLH1* promoter DNA in a bodily fluid from the subject;

wherein presence of the methylated *hMLH1* promoter DNA in the bodily fluid indicates a poor prognosis or the subject is a candidate for more aggressive therapy.

29. A primer set for detecting the presence of a cancer associated with methylation of *hMHL1* promoter DNA in a subject;

wherein the primer set is a first primer set and comprises a forward primer and the reverse primer for amplifying a bisulfite converted template in a CpG rich region of the Watson strand of *hMLH1* promoter DNA,

wherein said forward primer of the first primer set comprises a sequence which is substantially identical to a first target sequence in the bisulfite converted template in the Watson strand, said first target sequence being at least 8 nucleotides in length, said first target sequence being located at the 5' end of the bisulfite-converted template; and

wherein said reverse primer of set first primer set comprises a sequence which is identical to the reverse complement of a second target sequence in the bisulfite converted template in the Watson strand, said second target sequence being at least 8 nucleoitdes in length, said second target sequence being located at the 3'end of the bisulfite-converted template; or

wherein the primer set is a second primer set and comprises a forward primer and the reverse primer for amplifying a bisulfite converted template in a CpG rich region of the Crick strand of *hMLH1* promoter DNA,

wherein said forward primer of the second primer set comprises a sequence which is substantially identical to a first target sequence in the bisulfite converted template of the Crick strand, said first target sequence being at least 8 nucleotides in length, said first target sequence being located at the 5' end of the bisulfite-converted template; and

wherein said reverse primer of set second primer set comprises a sequence which is identical to the reverse complement of a second target sequence in the bisulfite converted template in the Crick strand, said second target sequence being at least 8 nucleoitdes in length, said second target sequence being located at the 3'end of the bisulfite-converted template

30. The primer set of claim 29 wherein each primer comprises a CpG dinucleotide.

31. The primer set of claim 29 wherein the forward primer, or the reverse primer, or both comprise at least two CpG dinucleotides.

32. The primer set of claim 29 wherein the forward primer comprises at least two thymines which correspond to two uracils that are immediately 5' of an A, C, or T base in the bisulfite-converted template sequence.

33. The primer set of claim 29 wherein the reverse primer comprises at least two adenines which are complementary to two uracils that are immediately 5' of an A, C, or T base in the bisulfite-converted template sequence

34. The primer set of claim 29 wherein the forward primer comprises at least 8 consecutive nucleotides in SEQ ID NO. 2 and the reverse primer comprises at least 8 consecutive nucleotides in SEQ ID NO. 3.

35. The primer set of claim 29 wherein the forward primer comprises at least 8 consecutive nucleotides in SEQ ID NO. 7 and the reverse primer comprises at least 8 consecutive nucleotide in SEQ ID NO. 8.

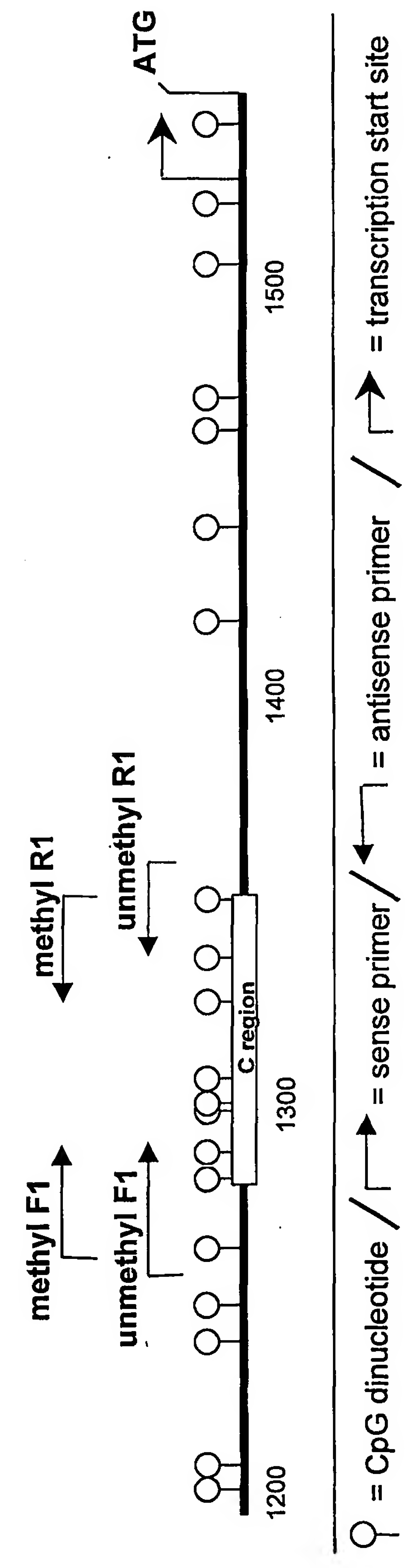

ATG
1500
1400
1300
1200
methyl R1
unmethyl R1
methyl F1
unmethyl F1
C region
= CpG dinucleotide / = sense primer / = antisense primer / = transcription start site

FIGURE 1

```
1    acgcttccat ttcttttctt agagagaata ttagttttaa agttttccct ctcctaagcc
61   aattgttcag acatcgtcca ttaacccttt caccatgctc tgattcagat tccgttagta
121  aagtctgaaa ttcaggtcgc ctaacgcctc tggtcttaat caaaactatc cattcggata
181  tggtatttat gatctgatta acatctacat cataggagct caacagttcc aagtgaagaa
241  atcctgaata atgaatgaca ttccaacagt tggggcttac ggccttttct ctcatgtatg
301  agacgagtcc tgagttctca ccaatcatcc tcaaagtatg ggtcgtggtc agtccaacca
361  ttctctgcaa gctaagccct gctgtctgca gggactctag gattgccgac atgagcgcac
421  caacactgaa atgatgagtc aggttgatta tggtcagaag atcttcttgc acctccaact
481  cagggcctac accgcggata aagaccagga ggtagttctc ataggccaca aaagcctggt
541  cgtccaaggc aagagaatag gctttaaagt ccctggctcg gttaaaaagc tggttgcgta
601  gattcctgtc aatgctcagg atcctctgcc ttgtgatatc tggagataag tcaacgcctt
661  gcaggacgct tacatgctcg ggcagtacct ctctcagcaa cacctccatg cactggtata
721  caaagtcccc ctcaccccag ccgcgaccct tcaaggccaa gaggcggcag agcccgaggc
781  ctgcacgagc agctctctct tcaggagtga aggaggccac gggcaagtcg ccctgacgca
841  gacgctccac cagggccgcg cgctcgccgt ccgccacata ccgctcgtag tattcgtgct
901  cagcctcgta gtggcgcctg acgtcgcgtt cgcgggtagc tacgatgagg cggcgacaga
961  ccaggcacag ggccccatcg ccctccggag gctccaccac caaataacgc tgggtccact
1021 cgggccggaa aactagagcc tcgtcgactt ccatcttgct tcttttgggc gtcatccaca
1081 ttctgcggga ggccacaaga gcagggccaa cgttagaaag gccgcaaggg gagaggagga
1141 gcctgagaag cgccaagcac ctcctccgct ctgcgccaga tcacctcagc agaggcacac
1201 aagcccggtt ccggcatctc tgctcctatt ggctggatat ttcgtattcc ccgagctcct
                                                                  5'-t

     aaaaatgaat taataggaag agtggatagt g-3    (SEQ. ID. NO. 5)
  5'-aaaaatgaat taataggaag agtggatagt g-3'   (SEQ. ID. NO. 4)
     5'-aacgaat taataggaag agcggatagc g-3'   (SEQ. ID. NO. 2)
1261 aaaaacgaac caataggaag agcggacagc gatctctaac gcgcaagcgc atatccttct


                    3'-aca aaatccctcc ctacttctct aa-5' (SEQ. ID. NO. 6)
          3'-ccc atcatcagca aaatccctcc ctgc-5'         (SEQ. ID. NO. 3)
1321 aggtagcggg cagtagccgc ttcagggagg gacgaagaga cccagcaacc cacagagttg


1381 agaaatttga ctggcattca agctgtccaa tcaatagctg ccgctgaagg gtggggctgg
1441 atggcgtaag ctacagctga aggaagaacg tgagcacgag gcactgaggt gattggctga
1501 aggcacttcc gttgagcatc tagacgtttc cttggctctt ctggcgccaa aatg 1554   (SEQ. ID. NO. 1)
```

FIGURE 2

```
1       acgcttccat ttctttttctt agagagaata ttagttttaa agttttccct ctcctaagcc
        tgcgaaggta aagaaaagaa tctctcttat aatcaaaatt tcaaaaggga gaggattcgg

61      aattgttcag acatcgtcca ttaacccttt caccatgctc tgattcagat tccgttagta
        ttaacaagtc tgtagcaggt aattgggaaa gtggtacgag actaagtcta aggcaatcat

121     aagtctgaaa ttcaggtcgc ctaacgcctc tggtcttaat caaaactatc cattcggata
        ttcagacttt aagtccagcg gattgcggag accagaatta gttttgatag gtaagcctat

181     tggtatttat gatctgatta acatctacat cataggagct caacagttcc aagtgaagaa
        accataaata ctagactaat tgtagatgta gtatcctcga gttgtcaagg ttcacttctt

241     atcctgaata atgaatgaca ttccaacagt tggggcttac ggcctttttct ctcatgtatg
        taggacttat tacttactgt aaggttgtca accccgaatg ccggaaaaga gagtacatac

301     agacgagtcc tgagttctca ccaatcatcc tcaaagtatg ggtcgtggtc agtccaacca
        tctgctcagg actcaagagt ggttagtagg agtttcatac ccagcaccag tcaggttggt

361     ttctctgcaa gctaagccct gctgtctgca gggactctag gattgccgac atgagcgcac
        aagagacgtt cgattcggga cgacagacgt ccctgagatc ctaacggctg tactcgcgtg

421     caacactgaa atgatgagtc aggttgatta tggtcagaag atcttcttgc acctccaact
        gttgtgactt tactactcag tccaactaat accagtcttc tagaagaacg tggaggttga

481     cagggcctac accgcggata aagaccagga ggtagttctc ataggccaca aaagcctggt
        gtcccggatg tggcgcctat ttctggtcct ccatcaagag tatccggtgt tttcggacca

541     cgtccaaggc aagagaatag gctttaaagt ccctggctcg gttaaaaagc tggttgcgta
        gcaggttccg ttctcttatc cgaaatttca gggaccgagc caatttttcg accaacgcat

601     gattcctgtc aatgctcagg atcctctgcc ttgtgatatc tggagataag tcaacgcctt
        ctaaggacag ttacgagtcc taggagacgg aacactatag acctctattc agttgcggaa

661     gcaggacgct tacatgctcg ggcagtacct ctctcagcaa cacctccatg cactggtata
        cgtcctgcga atgtacgagc ccgtcatgga gagagtcgtt gtggaggtac gtgaccatat

721     caaagtcccc ctcaccccag ccgcgaccct tcaaggccaa gaggcggcag agcccgaggc
        gtttcagggg gagtggggtc ggcgctggga agttccggtt ctccgccgtc tcgggctccg

781     ctgcacgagc agctctctct tcaggagtga aggaggccac gggcaagtcg ccctgacgca
        gacgtgctcg tcgagagaga agtcctcact tcctccggtg cccgttcagc gggactgcgt

841     gacgctccac cagggccgcg cgctcgccgt ccgccacata ccgctcgtag tattcgtgct
        ctgcgaggtg gtcccggcgc gcgagcggca ggcggtgtat ggcgagcatc ataagcacga

901     cagcctcgta gtggcgcctg acgtcgcgtt cgcgggtagc tacgatgagg cggcgacaga
        gtcggagcat caccgcggac tgcagcgcaa gcgcccatcg atgctactcc gccgctgtct

961     ccaggcacag ggccccatcg ccctccggag gctccaccac caaataacgc tgggtccact
        ggtccgtgtc ccggggtagc gggaggcctc cgaggtggtg gtttattgcg acccaggtga

1021    cgggccggaa aactagagcc tcgtcgactt ccatcttgct tcttttgggc gtcatccaca
        gcccggcctt ttgatctcgg agcagctgaa ggtagaacga agaaaacccg cagtaggtgt

1081    ttctgcggga ggccacaaga gcagggccaa cgttagaaag gccgcaaggg gagaggagga
        aagacgccct ccggtgttct cgtcccggtt gcaatctttc cggcgttccc ctctcctcct

1141    gcctgagaag cgccaagcac ctcctccgct ctgcgccaga tcacctcagc agaggcacac
        cggactcttc gcggttcgtg gaggaggcga gacgcggtct agtggagtcg tctccgtgtg

1201    aagcccggtt ccggcatctc tgctcctatt ggctggatat ttcgtattcc ccgagctcct
        ttcgggccaa ggccgtagag acgaggataa ccgacctata aagcataagg ggctcgagga

1261    aaaaacgaac caataggaag agcggacagc gatctctaac gcgcaagcgc atatccttct
        tttttgcttg gttatccttc tcgcctgtcg ctagagattg cgcgttcgcg tataggaaga

1321    aggtagcggg cagtagccgc ttcagggagg gacgaagaga cccagcaacc cacagagttg
        tccatcgccc gtcatcggcg aagtccctcc ctgcttctct gggtcgttgg gtgtctcaac

1381    agaaatttga ctggcattca agctgtccaa tcaatagctg ccgctgaagg gtggggctgg
        tctttaaact gaccgtaagt tcgacaggtt agttatcgac ggcgacttcc caccccgacc

1441    atggcgtaag ctacagctga aggaagaacg tgagcacgag gcactgaggt gattggctga
        taccgcattc gatgtcgact tccttcttgc actcgtgctc cgtgactcca ctaaccgact

1501    aggcacttcc gttgagcatc tagacgtttc cttggctctt ctggcgccaa aatg  1554 SEQ. ID. NO. 1
        tccgtgaagg caactcgtag atctgcaaag gaaccgagaa gaccgcggtt ttac
```

FIGURE 3

FIGURE 4

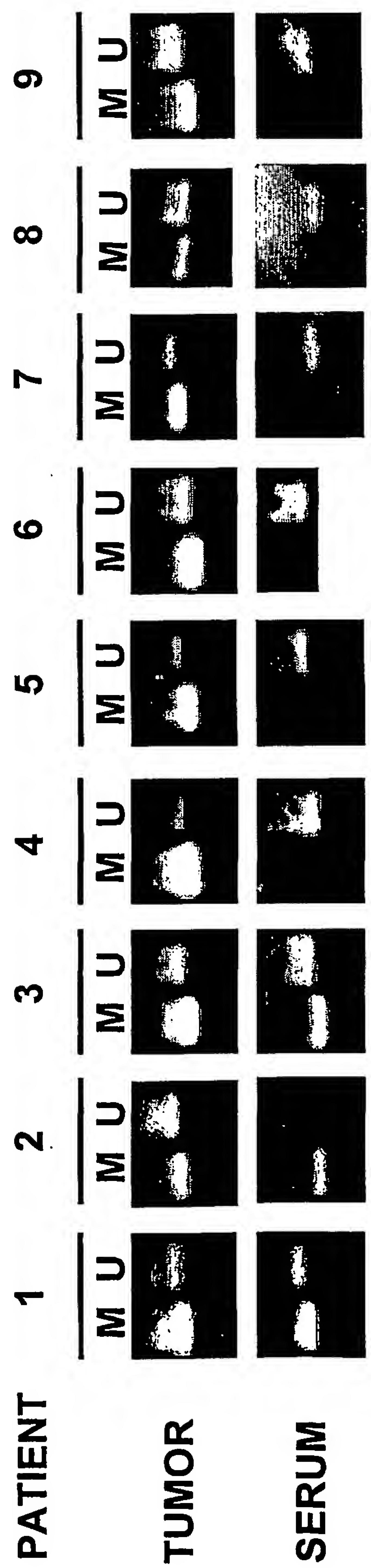


FIGURE 5

SEQUENCE LISTING

<110> Markowitz, Sanford
William, Grady

<120> Methods and Compositions for Detecting Cancers Associated w
ith Methylation of hMLH1 Promoter DNA

<130> 27708/04001

<140> 60/234,087
<141> 2000-09-20

<160> 8

<170> PatentIn version 3.0

<210> 1
<211> 1554
<212> DNA
<213> Human

<400> 1
acgcttccat ttcttttctt agagagaata ttagttttaa agttttccct ctcctaagcc
60

aattgttcag acatcgtcca ttaacccttt caccatgctc tgattcagat tccgttagta
120

aagtctgaaa ttcaggtcgc ctaacgcctc tggtcttaat caaaactatc cattcggata
180

tggtatttat gatctgatta acatctacat cataggagct caacagttcc aagtgaagaa
240

atcctgaata atgaatgaca ttccaacagt tggggcttac ggccttttct ctcatgtatg
300

agacgagtcc tgagttctca ccaatcatcc tcaaagtatg ggtcgtggtc agtccaacca
360

ttctctgcaa gctaagccct gctgtctgca gggactctag gattgccgac atgagcgcac
420

caacactgaa atgatgagtc aggttgatta tggtcagaag atcttcttgc acctccaact
480

cagggcctac accgcggata aagaccagga ggtagttctc ataggccaca aaagcctggt

540

cgtccaaggc aagagaatag gctttaaagt ccctggctcg gttaaaaagc tggttgcgta
600

gattcctgtc aatgctcagg atcctctgcc ttgtgatatc tggagataag tcaacgcctt
660

gcaggacgct tacatgctcg ggcagtacct ctctcagcaa cacctccatg cactggtata
720

caaagtcccc ctcaccccag ccgcgaccct tcaaggccaa gaggcggcag agcccgaggc
780

ctgcacgagc agctctctct tcaggagtga aggaggccac gggcaagtcg ccctgacgca
840

gacgctccac cagggccgcg cgctcgccgt ccgccacata ccgctcgtag tattcgtgct
900

cagcctcgta gtggcgcctg acgtcgcgtt cgcgggtagc tacgatgagg cggcgacaga
960

ccaggcacag ggccccatcg ccctccggag gctccaccac caaataacgc tgggtccact
1020

cgggccggaa aactagagcc tcgtcgactt ccatcttgct tcttttgggc gtcatccaca
1080

ttctgcggga ggccacaaga gcagggccaa cgttagaaag gccgcaaggg gagaggagga
1140

gcctgagaag cgccaagcac ctcctccgct ctgcgccaga tcacctcagc agaggcacac
1200

aagcccggtt ccggcatctc tgctcctatt ggctggatat ttcgtattcc ccgagctcct
1260

aaaaacgaac caataggaag agcggacagc gatctctaac gcgcaagcgc atatccttct
1320

aggtagcggg cagtagccgc ttcagggagg gacgaagaga cccagcaacc cacagagttg
1380

agaaatttga ctggcattca agctgtccaa tcaatagctg ccgctgaagg gtggggctgg
1440

atggcgtaag ctacagctga aggaagaacg tgagcacgag gcactgaggt gattggctga

1500

aggcacttcc gttgagcatc tagacgtttc cttggctctt ctggcgccaa aatg
1554

<210> 2
<211> 28
<212> DNA
<213> Synthetic primer

<400> 2
aacgaattaa taggaagagc ggatagcg
28

<210> 3
<211> 27
<212> DNA
<213> Synthetic primer

<400> 3
cgtccctccc taaaacgact actaccc
27

<210> 4
<211> 31
<212> DNA
<213> Synthetic primer

<400> 4
aaaaatgaat taataggaag agtggatagt g
31

<210> 5
<211> 32
<212> DNA
<213> Synthetic primer

<400> 5
taaaaatgaa ttaataggaa gagtggatag tg
32

<210> 6
<211> 25

<212> DNA
<213> Synthetic primer

<400> 6
aatctcttca tccctcccta aaaca
25

<210> 7
<211> 29
<212> DNA
<213> Synthetic primer

<400> 7
cgtttttttt tgaagcggtt attgtttgt
29

<210> 8
<211> 28
<212> DNA
<213> Synthetic primer

<400> 8
aacgaaccaa taaaaaaaac aaacaacg
28